

RECEIVED

JUN 13 2003

TECH CENTER 1600/2900



ENTERED

1600

RAW SEQUENCE LISTING

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:54

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

SEQUENCE LISTING

```

3 (1) GENERAL INFORMATION:
5   (i) APPLICANT: DOUGLAS SMITH
7   (ii) TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
8                               RELATING TO HELICOBACTER PYLORI FOR
9                               DIAGNOSTICS AND THERAPEUTICS
11  (iii) NUMBER OF SEQUENCES: 941
13  (iv) CORRESPONDENCE ADDRESS:
14      (A) ADDRESSEE: LAHIVE +ACY- COCKFIELD, LLP
15      (B) STREET: 28 State Street
16      (C) CITY: Boston
17      (D) STATE: Massachusetts
18      (E) COUNTRY: USA
19      (F) ZIP: 02109
21  (v) COMPUTER READABLE FORM:
22      (A) MEDIUM TYPE: Floppy disk
23      (B) COMPUTER: IBM PC compatible
24      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
25      (D) SOFTWARE: PatentIn Release +ACM-1.0, Version +ACM-1.25
27  (vi) CURRENT APPLICATION DATA:
C--> 28      (A) APPLICATION NUMBER: US/08/487,032C
C--> 29      (B) FILING DATE: 07-Jun-1995
31  (viii) ATTORNEY/AGENT INFORMATION:
32      (A) NAME: Mandragouras, Amy E.
33      (B) REGISTRATION NUMBER: 36,207
34      (C) REFERENCE/DOCKET NUMBER: GTN-001
36  (ix) TELECOMMUNICATION INFORMATION:
37      (A) TELEPHONE: (617)227-7400
38      (B) TELEFAX: (617)227-5941
41 (2) INFORMATION FOR SEQ ID NO: 1:
44   (i) SEQUENCE CHARACTERISTICS:
45      (A) LENGTH: 1527 base pairs
46      (B) TYPE: nucleic acid
47      (C) STRANDEDNESS: double
48      (D) TOPOLOGY: circular
50   (ii) MOLECULE TYPE: DNA (genomic)
52   (iii) HYPOTHETICAL: NO
54   (iv) ANTI-SENSE: NO
56   (vi) ORIGINAL SOURCE:
57      (A) ORGANISM: Helicobacter pylori
59   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
61 ATGTGTTCTC AGGAAATTTT ATCAAGCTTG CAAACCATTA TTGCCGAACA ATTTTCTATA      60
63 AATATCATCA CTCAGCTTGC TAATAAACTC ACACAAGTTA AAAATCTAAA TTTTTTTGAG      120

```

-RAW SEQUENCE LISTING

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:54

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

```

65 AATAAAGACC ATACTATCAA GCTTAACACT ATCCATAACG GACTGCACAT CCGCCCCCTA 180
67 AATTATGTCA GTAATCTTTT TTTCAATCTA CAACGCATTA TAGGGCTTAT CAGTCTGTTT 240
69 GGGATATTAT TTTCCATTAG TATTTATCTA CCCTTTATAA TGATTTTTC AACAGTGCCT 300
71 TGTATTCTCA TTTCCAACCA TATAGCAAAA AAACATAGTG CTTCCATAGA TAAACTTCAA 360
73 GACCAAAAAG AAAGCATGCA AAATTACTTA TACTCTGGAC TAGATAACCA AAAGAACAAG 420
75 GACAACCTAT TATTTAACTT CATGCTAAAT TTTCACCATA AATTTATTGA AACAAAAGAA 480
77 TTGTATCTCA ATAATTTTGT GAAAGTAGCC CAAAAAACT TAATATTTAC CATATATGCT 540
79 GATGTTTTAA TCACCACTCT AAGTATTGCA CTATTTTTC TAATGGTTTT TATTATCCTT 600
81 TCAAAATTAA TTGGTGTGGG AGCAATTGCT GGGTATATCC AAGCATTTAG CTCTACCCAA 660
83 CAACAACCTAC AAGATTTATC ATTTTATGGA AAGTGGTTTT TTGCTATCAA TAAATACTTT 720
85 GAAAATTATT TCTGTATTTT AGATTACAAA ATACCGAAAC CAGAAACACA AATCAAATTA 780
87 GAAGAAAAAA TCCATAGCAT TACATTTGAA AATATTAGTT TCTCTTATCC TAATTCAAAA 840
89 CTTATTTTTG AAAACTTTAA TCTCTCTTTA CACTCTAATA AAATTTATGC ATTAGTCGGC 900
91 AAGAATGCTA GCGGAAAAAG CACGCTGATT AATTTATTAT TAGGTTTTTA TACCCCAAAT 960
93 TCAGGTCAAA TTATCATTAA TAACAAATAC CCATTACAAG ACTTGGAAC AAATAGCTAC 1020
95 CATCAACAAA TGAGTGCCAT ATTTCAAGAT TTTTCTCTTT ATGCTGGGTA TAGCATTGAT 1080
97 GATAATCTTT TTATGCAAAA CAATATCACT AAAGAGCAAT TGAAGCAAAA AAGAGAAATA 1140
99 CTAAAATCTT TTGATGAGAA TTTTCAAAAT TGTCTTAATG ATTGCAACAA CACACTATTT 1200
101 GGAGCGCAAT ATAATGGGGT AGATTTTCT TTAGGTCAAA AGCAACGCAT AGCTACCATG 1260
103 AGAGCCTTTT TAAAACCAAG TAATTGCATT GTTTTAGATG AGCCAAGCAG CGCCATCGAT 1320
105 CCCATTATGG AAAAAGAGTT TTTAGATTTT ATTTTAAAAA AATCGCAATC TAAGATGGCT 1380
107 TTAATTATTA CACACCGCAT GAATAGTGTC AAGCAAGCTA ATGAAATTAT CGTGTTAGAT 1440
109 CAAGGCAAAC TAATAGAACA GGGCAACTTT GAAACCCTTA TGAAAAACA GGGATTATTT 1500
111 TGCGAATTGT TTTTGAAACA ACAATAC
114 (2) INFORMATION FOR SEQ ID NO: 2:
116 (i) SEQUENCE CHARACTERISTICS:
117 (A) LENGTH: 399 base pairs
118 (B) TYPE: nucleic acid
119 (C) STRANDEDNESS: double
120 (D) TOPOLOGY: circular
122 (ii) MOLECULE TYPE: DNA (genomic)
124 (iii) HYPOTHETICAL: NO
126 (iv) ANTI-SENSE: NO
128 (vi) ORIGINAL SOURCE:
129 (A) ORGANISM: Helicobacter pylori
131 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
133 ATGGAGCGCA AGACGCTCCA GAGCATTTTA TGTTTAATAA AAAAAGAAAT GATGAGACCA 60
135 AAAGGTATTC TAATGAATTG TTGCAGGSCT TGGAAACACC AGGTTCTTAA GCAAAGCACG 120
137 ACAGGTTTAT TGGTGTGAG CATTATCTCT TCTACAGCCC CCTTTATTGG TTTGTTTGGG 180
139 ACGGTAGTTG AAATTTTAGA AGCGTTTAAAC AATTTGGGCG CGTTAGGTCA AGCTTCTTTT 240
141 GGAGTGATCG CACCCATTAT TTCTAAGGCG CTTATCGCCA CCGCTGCAGG GATTTTAGCA 300
143 GCCATTCCAG CCTATTCTTT TACTTTGATC TTAAAGCGCA AGGTGTATGA TTTATCGGTT 360
145 TATGTGCAGA TGCAAGTGGA TATTTTGTCT TCTAAAAAA 399
148 (2) INFORMATION FOR SEQ ID NO: 3:
150 (i) SEQUENCE CHARACTERISTICS:
151 (A) LENGTH: 474 base pairs
152 (B) TYPE: nucleic acid
153 (C) STRANDEDNESS: double
154 (D) TOPOLOGY: circular

```

RAW SEQUENCE LISTING

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:54

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

```

156      (ii) MOLECULE TYPE: DNA (genomic)
158      (iii) HYPOTHETICAL: NO
160      (iv) ANTI-SENSE: NO
162      (vi) ORIGINAL SOURCE:
163          (A) ORGANISM: Helicobacter pylori
165      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
167 ATGCATGAAC GCATTGAAAG AGGTATTGGA AATAATGAAT GTAAAGAAAT TTTTGGCAAT      60
169 GAACTCAAAC AAAGAAAGAC AAAATTGATT GAAGACATAG AACGCGCGTT CAAAGAATGC      120
171 GAGGAACAAT TCCGTGGAAG TGTAGGAAAA AATATTGAAC AACTTGAAGA AAGAGTTAAA      180
173 GATTCTCTAG CGATTATAAA ACGCATCAAT AACCTTGGTC TTAATCCTAA TTCTAATTTT      240
175 AATATGGATA GCGGCATTGA TACAATAGGC TTATTTAGTT CAATAGGAGG TTTGGTGTTG      300
177 CTTCTATTGA CGCCTGTAGT AGGTGAGTTT GCGTTAATTG CAGGAGTGGG TTTAGCATTA      360
179 GTGGGGGTAG GTAAATCAAT ATGGAGTTTT TTTGATTGAG ATTATAAAAA ATCCCAACAA      420
181 AGAAAAGAAG TGGATAAGAA TTTACATCAA ATTTGCGAAA AATTGTGCAG GATG          474
184 (2) INFORMATION FOR SEQ ID NO: 4:
186      (i) SEQUENCE CHARACTERISTICS:
187          (A) LENGTH: 336 base pairs
188          (B) TYPE: nucleic acid
189          (C) STRANDEDNESS: double
190          (D) TOPOLOGY: circular
192      (ii) MOLECULE TYPE: DNA (genomic)
194      (iii) HYPOTHETICAL: NO
196      (iv) ANTI-SENSE: NO
198      (vi) ORIGINAL SOURCE:
199          (A) ORGANISM: Helicobacter pylori
201      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
204 ATGCCTGGCG TGTATCAAAT GAGTATAGAG CCTCTTTTAA AAGAATGCCA AGAATTAGTG      60
206 GGTTTAGGCA TAAAAGCCGT TTTATTGTTT GGCATTCCTA AACATAAGGA CGCTACAGGA      120
208 AGCCATGCGT TAAATAAGGA TCACATTGTC GCAAAGCTA CGAGAGAAAT TAAAAAACGA      180
210 TTTAAGGATT TGATCGTTAT AGCGGATTTG TGTTTTTGCG AATACACCGA CCATGGGCAT      240
212 TCGGGGATTT TAGAAAACGC TTCTGTGTCT AACGATAAAA CGCTAAAGAT TTTAAATCTT      300
214 CAAGGGCTTA TTTTGCTGAA AGCGGTGTGG ATATTC          336
217 (2) INFORMATION FOR SEQ ID NO: 5:
219      (i) SEQUENCE CHARACTERISTICS:
220          (A) LENGTH: 195 base pairs
221          (B) TYPE: nucleic acid
222          (C) STRANDEDNESS: double
223          (D) TOPOLOGY: circular
225      (ii) MOLECULE TYPE: DNA (genomic)
227      (iii) HYPOTHETICAL: NO
229      (iv) ANTI-SENSE: NO
231      (vi) ORIGINAL SOURCE:
232          (A) ORGANISM: Helicobacter pylori
234      (ix) FEATURE:
235          (A) NAME/KEY: misc+AF8-feature
236          (B) LOCATION: 1...195
237          (D) OTHER INFORMATION: /note+AD0- +ACI-FLAGELLAR HOOK-ASSOCIATED PROTEIN 1
238 HAP1+ACI-
240      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

-RAW SEQUENCE LISTING

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:54

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

```

242 GTGGAAAACA ACAAGAGTTT AAAGCATGCG AATGAGTTAA GGGATAAGCG AGATGAATTA      60
244 GAGTTTCATT TGCAGAGAGCT TTTCTGGGGGG AATGTTTTTA AAAGCAGCAT TAAAACCCAT      120
246 TCGCTCACAG ATAAAGACTC AGCGGACTTT GATGAGAGCT ATAACCTTAA TATCGGGCAT      180
248 GGGYTCAATA TSATA                                          195
251 (2) INFORMATION FOR SEQ ID NO: 6:
253     (i) SEQUENCE CHARACTERISTICS:
254         (A) LENGTH: 1857 base pairs
255         (B) TYPE: nucleic acid
256         (C) STRANDEDNESS: double
257         (D) TOPOLOGY: circular
259     (ii) MOLECULE TYPE: DNA (genomic)
261     (iii) HYPOTHETICAL: NO
263     (iv) ANTI-SENSE: NO
265     (vi) ORIGINAL SOURCE:
266         (A) ORGANISM: Helicobacter pylori
268     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
270 GTGTTTGTGG CAAGCAAACA AGCTGACGAA CAAAAAAGC TAGTTATAGA GCAAGAGGTT      60
272 CAAAAGCGGC AGTTTCAAAA AATAGAAGAA CTAAAGCAG ACATGCAAAA GGGTGTCAT      120
274 CCCTTTTTTA AAGTCTTGTT TGATGGGGGG AATAGGTTGT TTGGTTTCCC TGAAACTTTT      180
276 ATTTATTCTT CTATATTTAT ATTGTTTGTA ACAATTGTAT TATCTGTTAT TCTTTTTCAA      240
278 GCCTATGAAC CTGTTTTGAT TGTAGCGATT GTTATTGTGC TTGTAGCTCT TGGATTCAAG      300
280 AAAGATTACA GGCCTTATCA AAGAATGGAG CGAGCGATGA AATTTAAAAA ACCTTTTTTG      360
C--> 282 TTTAAGGGCG TGA AAAACAA AGCGTTCATG AGCATTTTTT CCATGAAGCC TAGTAAAGAA      420
284 ATGGCTAATG ACATCCACTT AAATCCAAAC AGAGAAGACA GGCTTGTGAG CGCTGCAAAC      480
286 TCCTATCTAG CGAATAACTA TGAATGTTTT TTAGATGATG GGGTGATCCT TACTAACAAC      540
288 TATTCTCTTT TAGGCACAAT CAAATTGGGG GGCATTGATT TTTTAACCAAC TTCCAAAAAA      600
290 GATCTCATAG AGTTACACGC TTCTATTTAT AGCGTTTTTA GGAATTTTGT TACCCCTGAA      660
292 TTCAAATTTT ATTTTCACAC TGTTAAAAAG AAAATCGTTA TTGATGAAAC CAATAGGGAT      720
294 TATGGTCTTA TTTTTTCTAA TGATTTTCATG CGAGCCTATA ATGAGAAGCA AAAGAGAGAA      780
296 AGTTTTTATG ATATTAGTTT TTATCTCACC ATAGAGCAAG ATTTATTAGA CACTCTCAAT      840
298 GAACCCGTTA TGAATAAAAA GCATTTTGCA GACAATAATT TTGAAGAGTT TCAAAGGATT      900
300 ATTAGAGCCA AGCTTGAAAA CTTCAAAGAT AGGATAGAGC TCATAGAAGA GCTACTGAGT      960
302 AAATACCACC CCACTAGATT AAAAGAATAC ACTAAAGATG GCATTATTTA CTCCAAACAA      1020
304 TGCGAATTTT ACAATTTTCT TGTGGGAATG AATGAAGCCC CTTTTATTTG CAACAGAAAA      1080
306 GACTTGTATC TCAAGGAAAA AATGCATGGT GGGGTGAAAG AAGTTTATTT TGCCAATAAG      1140
308 CATGGAAAAA TCTTAAATGA CGATTTGAGT GAAAAATATT TTAGCGCTAT TGAGATCAGT      1200
310 GAATACGCCC CTAAATCACA GAGCGATTTG TTTGATAAAA TCAACGCTCT AGACAGCGAA      1260
312 TTTATCTTTA TGCATGCTTA TTCGCCTAAA AACTCACAAG TTTTAAAGGA CAACTAGCT      1320
314 TTCACCTCTA GAAGGATTAT TATTAGTGGA GGCTCCAAAG AGCAAGGCAT GACTTTGGGT      1380
316 TGCTTGAGCG AATTAGTGGG TAATGGTGAT ATTACGCTAG GCAGTTATGG TAATCTTTA      1440
318 GTGCTGTTTG CTGATAGCTT TGAAAAAATG AAACAAAGCG TTAAGGAATG CGTCTCTAGT      1500
320 CTTAACGCTA AAGGTTTTTT AGCCAAACGCA GCGACTTTCT CTATGGAAAA TTACTTTTTT      1560
322 GCCAAACATT GCTCTTTTAT CACGCTTCCT TTTATTTTTG ATGTAACCTC TAACAATTTT      1620
324 GCTGATTTCA TAGCGATGAG AGCGATGAGT TTTGATGGCA AAGAAGACAA TAACGCTTGG      1680
326 GGCAATAGCG TGATGACGTT AAAAAGCGAG ATCAATTCGC CTTTTTATTT GAACTCCAC      1740
328 ATGCCCCTG ATTTTGTTT AGCTTCAGCA GGACACACTT TGATACTTGG CTCAACCGGT      1800
330 TCAGGTAAGA ACAGTGTTTA TGTCCATGAC TCTAAACGCT ATGGGGCAAT TTGCCTA      1857
333 (2) INFORMATION FOR SEQ ID NO: 7:
335     (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:54

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

```

336      (A) LENGTH: 330 base pairs
337      (B) TYPE: nucleic acid
338      (C) STRANDEDNESS: double
339      (D) TOPOLOGY: circular
341      (ii) MOLECULE TYPE: DNA (genomic)
343      (iii) HYPOTHETICAL: NO
345      (iv) ANTI-SENSE: NO
347      (vi) ORIGINAL SOURCE:
348          (A) ORGANISM: Helicobacter pylori
350      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
352 GTGAAAACAT CGTGTGTTGGT TACAATAGGG AGGATCCGGG GCGTTTTTAT CATTAAAGGCG      60
354 CAGTTGTTGC TTCGTGAGGG AGGTTTTATG AATTTTACCG CTTATAACAC GAAGACGCCA      120
356 GGGCATTGTC ATTTGTATGT GCATAAGGGG CATACGGAAT TAGGCGAGGG TGAAAGGCTG      180
358 ATTAAACTT TATCCATGAA ATTAGCGCAA GGGTGCCTA AAGAATGGAG GGTTCCTCCT      240
360 AGCAATGAAT GGCCTAAGGA ATTTAATATT TTAGCTTAC CTTATGAAGT GTTGCAGAAA      300
362 GAGCGCGGGA GCTCTGGGC GAAGCATTTA      330
365 (2) INFORMATION FOR SEQ ID NO: 8:
367      (i) SEQUENCE CHARACTERISTICS:
368          (A) LENGTH: 204 base pairs
369          (B) TYPE: nucleic acid
370          (C) STRANDEDNESS: double
371          (D) TOPOLOGY: circular
373      (ii) MOLECULE TYPE: DNA (genomic)
375      (iii) HYPOTHETICAL: NO
377      (iv) ANTI-SENSE: NO
379      (vi) ORIGINAL SOURCE:
380          (A) ORGANISM: Helicobacter pylori
382      (ix) FEATURE:
383          (A) NAME/KEY: misc+AF8-feature
384          (B) LOCATION: 1...204
385          (D) OTHER INFORMATION: /note+AD0- +ACI-H+-transporting ATP synthase
386 alpha chain homolog+ACI-
388      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
390 GTGGCTAAAG ACATCATCAG CGAGTCTCAA AACCTTTGCG CAAGAAAATT CCGCCGTTTG      60
392 TATGCGTTAT TGAAAGAAAA TGAAATGCTC ATTCGCATCG GATCTTATCA AATGGGGAAC      120
394 GATAAAGAGC TTGATGAAGC GATTAAGAAA AAGGCTCTAA TGGAGCAATT TTTAGTGCAA      180
396 GATGAAAACG CTTTACYAGC CTTT      204
399 (2) INFORMATION FOR SEQ ID NO: 9:
401      (i) SEQUENCE CHARACTERISTICS:
402          (A) LENGTH: 1338 base pairs
403          (B) TYPE: nucleic acid
404          (C) STRANDEDNESS: double
405          (D) TOPOLOGY: circular
407      (ii) MOLECULE TYPE: DNA (genomic)
409      (iii) HYPOTHETICAL: NO
411      (iv) ANTI-SENSE: NO
413      (vi) ORIGINAL SOURCE:
414          (A) ORGANISM: Helicobacter pylori
416      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/08/487,032C

DATE: 06/06/2003
TIME: 08:04:55

Input Set : D:\Seqlistcorr3.txt
Output Set: N:\CRF4\06062003\H487032C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:387; Xaa Pos.2,3,19
Seq#:389; Xaa Pos.10
Seq#:391; Xaa Pos.62,64
Seq#:392; Xaa Pos.32,194
Seq#:395; Xaa Pos.66
Seq#:399; Xaa Pos.256,292
Seq#:402; Xaa Pos.270,272
Seq#:403; Xaa Pos.130,163
Seq#:407; Xaa Pos.4
Seq#:412; Xaa Pos.59,61,66,70
Seq#:421; Xaa Pos.89
Seq#:422; Xaa Pos.37
Seq#:423; Xaa Pos.136
Seq#:427; Xaa Pos.203
Seq#:432; Xaa Pos.52,53
Seq#:433; Xaa Pos.34,67,70,72,75,76
Seq#:437; Xaa Pos.8
Seq#:449; Xaa Pos.50
Seq#:454; Xaa Pos.124,174
Seq#:458; Xaa Pos.51
Seq#:461; Xaa Pos.175
Seq#:467; Xaa Pos.5
Seq#:490; Xaa Pos.123,126,135,136,141,144,148,152,164,167
Seq#:494; Xaa Pos.12,20,41,52,53,64,65,66
Seq#:499; Xaa Pos.14,34,37,39,40,41,42,43,44,46,48,49,63,71
Seq#:502; Xaa Pos.169,172,175,176,179,182,184
Seq#:508; Xaa Pos.30
Seq#:509; Xaa Pos.48,117
Seq#:519; Xaa Pos.70,72,109
Seq#:523; Xaa Pos.55,57,65
Seq#:524; Xaa Pos.97,98,102
Seq#:527; Xaa Pos.2,21,29
Seq#:530; Xaa Pos.13
Seq#:533; Xaa Pos.56,57,99,102
Seq#:539; Xaa Pos.99,112,113,130
Seq#:544; Xaa Pos.2,12
Seq#:551; Xaa Pos.9,34,42
Seq#:556; Xaa Pos.116
Seq#:558; Xaa Pos.7,13,17
Seq#:560; Xaa Pos.27,28
Seq#:565; Xaa Pos.9,47
Seq#:570; Xaa Pos.14,34,55,58,65
Seq#:573; Xaa Pos.65,72
Seq#:578; Xaa Pos.3,4,10,11,27

·RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/08/487,032C

DATE: 06/06/2003
TIME: 08:04:55

Input Set : D:\Seqlistcorr3.txt
Output Set: N:\CRF4\06062003\H487032C.raw

Seq#:582; Xaa Pos.2,101,142,149,152,157,161
Seq#:587; Xaa Pos.25
Seq#:589; Xaa Pos.18,22,44
Seq#:598; Xaa Pos.14,138,154
Seq#:602; Xaa Pos.4
Seq#:608; Xaa Pos.27
Seq#:609; Xaa Pos.48
Seq#:617; Xaa Pos.15,132,133,134,139,140,141,143

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:5; Line(s) 237
Seq#:48; Line(s) 1849
Seq#:53; Line(s) 2030
Seq#:74; Line(s) 2822
Seq#:78; Line(s) 3132
Seq#:93; Line(s) 3759
Seq#:159; Line(s) 6410
Seq#:161; Line(s) 6491
Seq#:170; Line(s) 6843
Seq#:215; Line(s) 8712
Seq#:218; Line(s) 8827
Seq#:233; Line(s) 9402
Seq#:245; Line(s) 9833
Seq#:256; Line(s) 10267
Seq#:258; Line(s) 10347
Seq#:274; Line(s) 10999
Seq#:279; Line(s) 11196
Seq#:285; Line(s) 11430
Seq#:295; Line(s) 11757
Seq#:304; Line(s) 12093
Seq#:350; Line(s) 13856
Seq#:388; Line(s) 15487
Seq#:393; Line(s) 15716
Seq#:410; Line(s) 16715,16716
Seq#:424; Line(s) 17346
Seq#:437; Line(s) 17842
Seq#:448; Line(s) 18408
Seq#:456; Line(s) 18779
Seq#:461; Line(s) 19020
Seq#:473; Line(s) 19581
Seq#:483; Line(s) 20031
Seq#:491; Line(s) 20717
Seq#:492; Line(s) 20758
Seq#:496; Line(s) 21033
Seq#:512; Line(s) 21837
Seq#:523; Line(s) 22390

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/08/487,032C

DATE: 06/06/2003
TIME: 08:04:55

Input Set : D:\Seqlistcorr3.txt
Output Set: N:\CRF4\06062003\H487032C.raw

Seq#:530; Line(s) 22651
Seq#:548; Line(s) 23642
Seq#:576; Line(s) 24997
Seq#:590; Line(s) 25712
Seq#:596; Line(s) 26098
Seq#:609; Line(s) 26903
Seq#:610; Line(s) 26983
Seq#:611; Line(s) 27018
Seq#:616; Line(s) 27255
Seq#:638; Line(s) 28497
Seq#:644; Line(s) 28734
Seq#:651; Line(s) 29101
Seq#:652; Line(s) 29168
Seq#:658; Line(s) 29529
Seq#:666; Line(s) 30109
Seq#:686; Line(s) 31138

VERIFICATION SUMMARY

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:55

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:282 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=6
L:2866 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=74
L:6688 M:111 C: (47) String data converted to upper case,
L:9501 M:111 C: (47) String data converted to upper case,
L:14031 M:111 C: (47) String data converted to upper case,
L:15430 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:387 after pos.:0
M:341 Repeated in SeqNo=387
L:15535 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:389 after pos.:0
L:15636 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:391 after pos.:48
L:15662 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:392 after pos.:16
M:341 Repeated in SeqNo=392
L:15911 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:395 after pos.:64
L:16152 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:399 after pos.:240
M:341 Repeated in SeqNo=399
L:16337 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:402 after pos.:256
L:16382 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:403 after pos.:128
M:341 Repeated in SeqNo=403
L:16587 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:407 after pos.:0
L:16823 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:412 after pos.:48
M:341 Repeated in SeqNo=412
L:17240 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:421 after pos.:80
L:17269 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:422 after pos.:32
L:17325 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:423 after pos.:128
L:17488 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:427 after pos.:192
L:17665 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:432 after pos.:48
L:17697 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:433 after pos.:32
M:341 Repeated in SeqNo=433
L:17846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:437 after pos.:0
L:18453 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:449 after pos.:48
L:18720 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:454 after pos.:112
M:341 Repeated in SeqNo=454
L:18868 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:458 after pos.:48
L:19055 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:461 after pos.:160
L:19317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:467 after pos.:0
L:20677 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:490 after pos.:112
M:341 Repeated in SeqNo=490
L:20893 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:494 after pos.:0
M:341 Repeated in SeqNo=494
L:21189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:499 after pos.:0
M:341 Repeated in SeqNo=499
L:21343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:502 after pos.:160
M:341 Repeated in SeqNo=502
L:21620 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:508 after pos.:16
L:21660 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:509 after pos.:32

-VERIFICATION SUMMARY

DATE: 06/06/2003

PATENT APPLICATION: US/08/487,032C

TIME: 08:04:55

Input Set : D:\Seqlistcorr3.txt

Output Set: N:\CRF4\06062003\H487032C.raw

M:341 Repeated in SeqNo=509
L:22236 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:519 after pos.:64
M:341 Repeated in SeqNo=519
L:22403 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:523 after pos.:48
M:341 Repeated in SeqNo=523
L:22450 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:524 after pos.:96
L:22552 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:527 after pos.:0
M:341 Repeated in SeqNo=527
L:22655 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:530 after pos.:0
L:22989 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:533 after pos.:48
M:341 Repeated in SeqNo=533
L:23222 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:539 after pos.:96
L:41959 M:220 C: Keyword misspelled or invalid format, [(D) TOPOLOGY:]
L:41959 M:238 W: Alpha Fields not Ordered, Reordered [(D) TOPOLOGY:] of (2)(i)
L:41959 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=906